DOCUMENT RESUME

ED 190 615 TM 800 409

AUTHOR Ohara, Takeshi: Purcell, Thomas D.
TITLE Do Student Reported Achievements Have Validity as Measures of Instructional Efficacy? A Study of the Relationship between Instructional Context and Student Reported Achievements.
PUB DATE Apr 80
NOTE 12p.: Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April 7-11, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement: Classroom Environment: College Science: Correlation: Course Content: *Course Evaluation: Difficulty Level: Factor Structure: Higher Education: Questionnaires: *Self Evaluation (Individuals): *Student Evaluation of Teacher Performance: Teaching Styles: *Test Validity
IDENTIFIERS Instructor Course Evaluation Questionnaire

ABSTRACT

The relationship between student reported achievements (SRA) and student ratings of their corresponding course and teacher was investigated, from responses to the Instructor and Course Evaluation Questionnaire. The SRA consisted of 10 commonly accepted concepts of cognitive and individual development; their average was used as a measure of instructional efficacy. The ratings were organized by four factors. (1) personal-interpersonal relationship between student and teacher: (2) course structure: (3) course quality: and (4) course difficulty. There were ten groups representing a combination of grade level, freshman to graduate, and a science versus non-science dichotomy. Correlations between SRA and the first three factors were significant, supporting the thesis that self-reported achievement may be substituted for actual achievement. After regression, course quality and structure were the most important predictors of instructional efficacy for eight of the ten groups. In the science freshman group, the personal-interpersonal factor was most important: no single factor contributed to instructional efficacy of science sophomores. Contrary to expectation, course difficulty was not a strong predictor for any group, implying that it is irrelevant to SRA if course structure and quality are adequate. (CP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED190615

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Paper Presented to a Meeting of

American Educational Research Association

Boston

April, 1980

Do Student Reported Achievements Have Validity as Measures of Instructional Efficacy? A Study of the Relationship Between Instructional Context and Student Reported Achievements.

Takeshi Ohara and Thomas D. Purcell

Southern Illinois University at Carbondale

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
T. Ohara
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 800409

INTRODUCTION

In situations in which a uniform instrument for instructional evaluation based on student ratings is used on a campus wide basis to determine instructors' teaching efficacy, then considerations of class characteristics and course subject area become critical (Centra, 1976). Typically in such questionnaires, some items are more relevant to a particular class for description and evaluation of instructional activities than for other classes. Students may give more careful thought in responding to the relevant items than for the irrelevant areas. One of the solutions to reduce the item irrelevancy problem is to allow instructors to choose from an item pool those specific items relevant to their courses. If this solution is chosen, however, a new problem results. Since each instructional situation is unique, it becomes difficult to accumulate a sufficiently large sample to develop normative data.

Another possible solution to the item irrelevancy problem is to group courses on the basis of course similarity (e.g., course level and subject area). Presumably, the same items would be appropriate for courses that are highly similar.

In any case, in calculating norms, it becomes important to know how the relationship between student ratings and student achievements is affected by course characteristics. If the relationship between then remains constant while course characteristics vary, we may ignore differences in course characteristics in calculating norms.

The purpose of the present study was to investigate how the relationship between student ratings of instruction and student reported achievements (SRA) varied across course level and subject area (science vs. non-science course orientation).

METHOD--Variables

The Instructor and Course Evaluation (ICE) questionnaire has been used on a voluntary basis for some time as a campus wide teacher and course evaluation instrument at Southern Illinois University at Carbondale. This instrument was designed to measure instructional characteristics, class characteristics and student reported achievements. The data used in the present study were taken from the results of student evaluations of 2,910 courses at Southern Illinois University at Carbondale, during 1977-1979. The courses included in this study came from virtually every department on campus.

Student Rating Variables

It has been reported that instructional efficacy may be related to the degree of student achievements (American Psychological Association 1966, Centra 1977, Frey 1975). This relationship underscores the thesis that the more a student achieves in a course, the better rating a student will give to an instructor. In the ICE, there are ten SRA items. They are:

Gaining factual knowledge.
Learning fundamental principles, generalizations, or theories.
Learning to apply course material to improve rational thinking.
Developing specific competencies needed by professionals in this field.
Learning how professionals in this field gain new knowledge.
Developing creative capacities.
Developing a sense of personal responsibility.
Gaining a broader appreciation of intellectual-cultural activity.
Developing skill in expressing myself orally or in writing.
Discovering the implications of the course material for understanding myself.

These items are rated in terms of a five point scale, each of which is termed as exceptional progress (5), very good progress (4), good progress (3), little progress (2), and no progress (1). The average of these ten items was used as a measure of instructional efficacy based on a content similarity between the items and commonly accepted educational concepts of achievement.

With regard to instructional characteristics, a factor analytic study (Miller, 1978) was conducted on 40 evaluations items. Four major factors were found termed as: personal-interpersonal, course structure, course difficulty, and course quality.

The personal-interpersonal factor correlated with items reflecting interpersonal communication skills, interest, and sensitivity to students' needs. The items which loaded heavily were:

Knew if students understood her/him.
Spoke understandably.
Answered impromptu questions satisfactorily.
Showed an interest in the course.
Accepted criticism and suggestions.
Showed an interest in students.
Was available outside of class.
Encouraged student participation.

12.52 percent of the total variance was accounted for by this factor.

Course structure was the second factor. This factor related to items assessing instructional skills in managing the course, in testing and grading, and in presenting concepts of the subject area. The following items were heavily loaded in this factor:

Made clear assignments.
Prepared for class.
Set clear standards for grading.
Graded fairly.
Gave several examples to explain complex ideas.
Increased your appreciation for the subject.
Organized and presented subject matter well.
Specified objectives of the course.
Achieved the specified objectives of the course.
Explained the subject clearly.
Was enthusiastic about the subject.
In general, taught the class effectively.
The course was well organized.
The tests covered the course material well.
Audio-visuals could be used more effectively.

This factor shared 21.04 percent of the total variance.

Course quality was the third factor. In general, this factor reflected course relevance and interest to the student, uniqueness or redundancy of

course material for the student and satisfaction with textual material (Miller, 1978). The following items were included on this factor:

This course was a good learning experience.
The content of this course was good.
I had trouble paying attention in class.
This course was very interesting.
This course was one of the best I have taken.
This course was a waste of time.
The textbook was good.
This course should be taught in some other way.
I covered this material in other courses.
This course should continue to be offered.
Generally, the course was good.

The shared variance of this factor was 20.23 percent of the total variance.

The fourth factor related to the degree of course difficulty. The items which loaded on this factor described the perceived difficulty of course material and ideas. They were:

There should be additional prerequisites.
There should be fewer prerequisites.
The amount of required work was appropriate.
The course material was too difficult.
The reading assignments were hard to understand.
I was often confused.

This factor shared 9.27 percent of the total variance.

The class characteristics section of the ICE included items describing the number of the students taking the class as required or elective, the number of study hours per day reported, the student's college, the number of students at each class level, the number of students of each sex, the number of students with each expected grade, and the number of transfer students. In the present study, initial attention was on the course characteristics (controlled by departmental practices) rather than the class characteristics (controlled by student choice behavior).

The course characteristic variables concerned here were course level and science vs. non-science subject orientation. Course level ranged from freshman to graduate and may represent the degree of complexities in curriculum and instruction, whereas the science vs. non-science dichotomy may

reveal differences in other course characteristics. It has been reported that items assessing course structure seem to be more important in science oriented courses than items relevant to interpersonal relations (Pohlmann, 1978). Thus, the combination of course level and the science vs. non-science dichotomy was chosen as a means of detecting differences in instructional characteristics. In combination, the two considerations defined ten groups of courses (5 course levels and 2 subject area types).

Statistical Analysis

The first analysis consisted of calculating a descriptive statistic (i.e., correlation coefficient) between the SRA items and the four ICE factors (personal-interpersonal, course structure and quality, and course difficulty) for each of the ten groups. The overall average of the ten SRA items was then regressed against each of the four ICE factors for each of the ten groups. All statistical analyses were conducted using SPSS subroutines such as PEARSON CORR and REGRESSION.

Results and Discussion

Table 1 contains the average of the zero order correlations between the ten SRA items and the four ICE factors for each of the ten groups. In general, SRA was more highly correlated with personal-interpersonal relations, course structure and quality than with course difficulty. Course difficulty, thus may not be as critical in reflecting instructional efficacy as the other three factors. The fact that the correlations for the three factors which reflect teaching context are significant adds strength to the thesis that the ten self-report achievement items may be used as substitutes for actual achievement. These results add "construct validity" support to the "content validity" assumptions regarding the nature of SRA items presented earlier. If additional research supports this interpretation, it may be possible to conduct research in this area with greater convenience.

It is generally much easier to obtain student responses to SRA-type items on an anonymous questionnaire than to obtain unbiased teacher-course-evaluation data when student identity must be known and matched to course performance data.

Table 2 presents the results from the stepwise regression analyses, creating the average of the ten SRA items as a single criterion variable and the four instructional characteristics as independent variables. The regression equation was used to examine the relative importance of each independent variable. The way the regression was used in this study needs some clarification with regard to multicollinearity in the independent variables and shrinkage of the multiple correlation. In terms of multicollinearity problem, the range of the correlation coefficients among the independent variables were from 0.20 to 0.70. Thus, extreme collinearity seems not to exist in the independent variables (Nie et. al., 1975).

Comparing the regression weights from several samples with different sample size, the shrinkage or the overestimation of the multiple correlation becomes critical (Kerlinger, 1973). The degree of the overestimation of the multiple correlation is affected, among other things, by the ratio of the number of independent variables to the size of the sample. In the current study, the sample size varied across the samples. Note the R^2 from group 9 (R^2=0.804, n=40) and group 5 (R^2=0.448, n=499) in Table 2. It appears that the overestimation of the multiple correlation may not be a critical problem in the current study. Since our main research interest was the unique contribution of each independent variable in describing instructional efficacy (SRA) in each sample, the variation in the magnitude of R^2 across the 10 samples is of relatively small consequence.

The results obtained from the regression of eight of the 10 groups indicated that course quality and course structure were the most important

instructional traits in predicting SRA. Groups of science freshman courses and science sophomore courses were different from the other groups. In the science freshman group, the personal-interpersonal factor was the most important trait influencing instructional efficacy, whereas in the science sophomore group no single factor contributed in a practical sense to instructional efficacy.

An unexpected result was that course difficulty factor was not a strong predictor of SRA in any group. This implies that if instruction is presented in a well organized manner and course quality adequate then course difficulty will be irrelevant to student reported achievements.

REFERENCES

American Psychological Assoication. Standards for Educational and Psychological Tests and Manuals. Washington, D.C.: American Psychological Association, 1966.

Centra, J. A., and Creech, F. R. The Relationship Between Student, Teachers, and Course Characteristics and Student Ratings of Teacher Effectiveness. Report No. PR-76-1. Princeton, N. J.: Educational Testing Service, 1976.

Centra, J. A..Using Student Assessments to Improve Performance and Vitality. New Directions for Institutional Research, 20, 1978.

Frey, P. W., Leonard, D. W.,& Beatty, W. W. Student Ratings of Instruction: Validation Research. American Educational Research Journal, 1975, 12, 435-447.

Miller, W. G.. Guidelines for Interpreting Results of the Instructional Evaluation, Southern Illinois University at Carbondale, 1978.

Nie, N. H. et. al..Statistical Package for the Social Sciences, New York.: McGraw-Hill, INc.,1975.

Pohlmann, J. T.. A Multivariate Analysis of Selected Class Characteristic and Student Ratings of Instruction. Multivariate Behavioral Research, 10, 1975.

Table 1

The Average of the Zero Order Correlation Coefficients Between the ten SRA Items and the Four Factors in ten Groups

		Inter-personal			Structure			Quality			Difficulty		
GROUP	n	ave	max	min	ave	max	min	ave	max	min	ave	max	min
Freshman,science	106	58	70	40	56	78	42	55	72	31	41	45	29
Sophomore,science	67	65	74	55	68	81	59	58	78	34	45	59	25
Junior,science	111	67	77	51	68	81	41	64	74	37	42	47	33
Senior,science	170	61	73	51	60	72	39	60	75	42	45	55	32
Graduate,science	35	63	78	28	66	82	33	60	84	22	33	47	20
Freshman,nonscience	498	58	65	42	58	73	30	56	67	26	32	43	24
Sophomore,nonscience	216	68	75	61	75	80	73	55	63	47	44	53	34
Junior,nonscience	410	68	72	64	69	77	55	72	80	59	46	53	38
Senior,nonscience	390	64	69	58	67	73	53	65	73	49	41	48	37
Graduate,nonscience	173	65	72	59	63	72	49	62	67	51	28	44	17

Table 2

The R^2 and the Standardized weights from the ten Regression Analyses, Treating the Average of the SRA Items as a single Criterion Variable and the Four Factors as Independent Variables.

		Personal Interpersonal		Course Structure		Course Quality		Course Difficulty	
		R^2	beta	R^2	beta	R^2	beta	R^2	beta
Freshman,science	107	.320	.256	.001	.077	.057	.289	.001	.081
Sophomore,science	70	.024	-.479	.009	.629	.007	-.129	.067	.245
Junior,science	113	.019	.150	.001	.078	.335	.492	.019	-.224
Senior,science	173	.004	-.179	.415	.628	.018	.190	.005	.089
Graduate,science	40	.002	.099	.035	-.273	.755	.970	.012	-.140
Freshman,nonscience	499	.011	.243	.374	.183	.063	.318	.000	-.006
Sophomore,nonscience	216	.002	-.095	.680	.928	-	-	.000	-.026
Junior,nonscience	415	.039	.255	.002	.120	.657	.559	.003	-.076
Senior,nonscience	397	.007	.197	.28	.186	.025	.251	.001	-.038
Graduate,nonscience	176	.059	.246	.008	.155	.320	.327	.006	-.086
All groups	2910	.006	.170	.034	.138	.376	.468	.008	-.123